

---

---

## **XVI Congreso Nacional de Tecnologías de la Información Geográfica 25, 26 y 27 de Junio de 2014. Alicante.**

---

---

# **Análisis multivariante de imagen de alta resolución para la identificación y caracterización de paisajes en ambiente fluvial**

**Dinameres Aparecida Antunes<sup>1</sup>, Carla Eva Prichoa<sup>2\*</sup>, Selma Regina Aranha  
Ribeiro<sup>1</sup>**

<sup>1</sup>*Universidade Estadual de Ponta Grossa*

<sup>2</sup>*Universidad Autónoma de Madrid  
Ciudad Universitaria Cantoblanco*

---

### **Resumen**

A través del Análisis Orientado a Objeto aplicado a una escena de satélite de alta resolución espacial originó un conjunto de 39 variables con características espaciales, espectrales y texturales. La escena pertenece a la cuenca del río Pitangui situada en el centro-este del estado de Paraná, Brasil. Se llevó a cabo dos Análisis de Componentes Principales con las 39 variables que resultaron respectivamente 22 y 14 variables. Se aplicó el análisis de conglomerados, que organizó en subgrupos variables de mayor similitud con la unidad de paisaje. El resultado muestra que las variables espectrales fueran las más significativas, a continuación las espaciales y finalmente las texturales. La reflectância mayor de los niveles digitales se refiere tanto a la vegetación predominante en el área, así como por los usos cerca de la zona ribera.

Palabras clave: Análisis Cluster, Análisis Orientado a Objetos, Análisis de Componentes Principales, Pléyades

---

---

\* E-mail: [cprichoa@yahoo.com.br](mailto:cprichoa@yahoo.com.br)

## 1. Introducción y objetivos propuestos

El objetivo de este trabajo es la caracterización de la paisagem mediante Análise Orientada a Objetos e Análise Multivariada, en una imagen Pléiades del paisaje fluvial del uno segmento do río Pitangui, Paraná, Brasil, y que presenta diferencias estructurales y topográficas regionales.

El Análisis Orientado a Objetos (AOO), el más indicado ya que, en la identificación de las clases, considera tanto la respuesta espectral de los objetos como sus variantes geométricas, topológicas y de textura (Muller, 1997; Antunes, 2003). Divide la imagen en partes/objetos que la constituyen, y el proceso finaliza cuando los objetos de interés en la aplicación se encuentran aislados. (Venturieri & Santos 1998). El nivel de abstracción deja de ser un píxel para convertirse en un objeto/región.

Se utilizó el algoritmo de segmentación FLSA (Full Lambda Schedule Algorithm), responsable de la fusión de los segmentos adyacentes, basándose en la combinación de información espectral, espacial y de textura, dando como resultado la agrupación de píxeles en forma de región, que deben estar espacialmente próximos, poseer características espectrales similares e informar sobre su descriptor (espectral, espacial y de textura). (Robinson, Redding and Crisp; 2002).

El Análisis de Componentes Principales (ACP) se emplea cuando existe un número considerable de datos con la intención de reducir la dimensión de los mismos en una cantidad aceptable y coherente, sin pérdida significativa de información (Bouroche; Saprota, 1982, Mather, 1999, Schowengerdt, 1997).

Para fijar los componentes principales, es necesario calcular la matriz de varianzas-covarianzas ( $\Sigma$ ), o la matriz de correlación ( $R$ ), encontrar los autovalores y los autovectores y, finalmente, escribir las combinaciones lineales resultantes que serán las nuevas variables, denominadas de componentes principales. Cada componente principal es una combinación lineal de todas las variables originales, independientes entre sí y estimadas con el propósito de mantenerse, en orden de estimación y en términos de variación total, la cual está incluida en los datos iniciales. Su interpretación se realiza mediante el grado de importancia, o mediante la influencia que cada variable tiene sobre cada componente, siendo que esta importancia viene dada por la correlación entre cada variable  $X_j$  y el componente  $Y_i$  que esté siendo interpretado. (Regazzi, 2001).

Se propone la técnica de Agrupaciones junto con el Análisis de Componentes Principales como forma de interpretación de variables procedentes de métodos automáticos y/o semiautomáticos como el Análisis Orientado a Objetos o de segmentación Full Lambda Schedule Algorithm - FLSA, ya que puede agrupar a aquellas que tengan características comunes, sin perder información del conjunto en estudio.

El Análisis de Agrupaciones es utilizado es un método que posibilita la interpretación individual de cada grupo y la relación que este grupo posee con el resto (Johnson; Wichern, 2002). Su aplicación engloba una gran variedad de técnicas y algoritmos (Vicini; Souza 2005) con el objetivo de encontrar y separar cada objeto en grupos similares. (Johnson & Wichern, 2002, Regazzi, 2001, Vicini; Souza 2005). Se aplica en análisis de datos ambientales, ya que posibilita un análisis simultáneo de los datos multidimensionales, compuesto por diversas variables, siendo su estructura conocida en la geometría como el hiperespacio o el espacio con muchas dimensiones. (Silva et al. 2007).

Silva et al. (2006) aplicaron el análisis integrado a datos de la naturaleza con el objetivo de encontrar homogeneidad en las unidades de paisaje. Mediante la aplicación del análisis multivariante detectó zonas

ambientales homogéneas obtenidas en función del medio biofísico, permitiendo la creación de un mapa de estas zonas y su estado de conservación.

Para emplear el Análisis de Agrupaciones en un conjunto de datos, es esencial la elección de un coeficiente que cuantifique el grado de semejanza entre dos objetos. Ese coeficiente puede ser dividido en categorías según la estimación de la medida de similitud, o diferencia, entre los individuos agrupados. Para esta investigación se optó por la distancia euclidiana, considerada la más adecuada, puesto que puede generar diferentes agrupaciones (Johnson & Wichnern, 2002, Mingoti, 2005).

## 2. Metodología de Trabajo

Se utilizó la imagen del satélite Pléiades-1A del 28 de enero de 2013, con las bandas multiespectrales: tres bandas del visible y una del infrarrojo cercano, todas con resolución espacial de 2 metros, espectral de 12 bits y temporal de 26 días. Se empleó la imagen (Figura 1) - en la composición coloreada, color verdadero - la cual destaca la morfología del río Pitangui, la vegetación de ribera y los usos de la agricultura en las zonas próximas al río.



Figura 1 – Imagen Pléiades utilizada.

La longitud de onda de cada banda multiespectral de la imagen Pléiades utilizada puede ser vista en la Tabla 1.

Tabla 1 – Características de las bandas multispectrales de la imagen Pléiades

Bandas espectrales		Longitud de onda ( $\mu$ )
1	Visible (verde)	0,50 - 0,62
2	Visible (rojo)	0,59 - 0,71
3	Visible (azul)	0,43 - 0,55
4	Infrarrojo cercano	0,74 - 0,94

Se aplicó en la imagen la técnica del Análisis Orientado a Objetos, el cual generó los vectores, y el banco de datos relacional conteniendo descriptores espaciales, espectrales y de textura de cada segmento creado. Este método utiliza propiedades básicas de los valores de los niveles digitales, así como las propiedades espaciales del perfil de los grupos formados tras cada repetición, separando aquellos con propiedades similares y espacialmente próximas y aislando los objetos de interés (PRICHOA, 2012).

Tras estandarizar las variables se realizaron pruebas por el Análisis de Componentes Principales, con el objetivo de reducir la dimensionalidad de los datos, de forma comitantemente con el Análisis de Agrupaciones (AA). Se realizaron dos procesamientos ACP y dos AA. En la ACP 1 hubo disminución de la dimensionalidad de los datos (descriptores) 36 para 22 y, en la ACP 2, 22 para 14. El criterio para la exclusión de las variables fue el grado de éstos, mediante análisis de los gráficos de los 5 componentes en el primer ACP y 3 componentes en el segundo ACP. Los descriptores de textura, espectrales y espaciales, que en los gráficos (CP1 - CP2, CP1 - CP3 y CP4 - CP5) quedaron aisladas y que tuvieron la mayor distancia en el AA fueron eliminados.

La última etapa fue la aplicación del Análisis de Agrupaciones que posibilitó la generación de dendrogramas, los cuales agruparon las variables con similitud permitiendo caracterizar las unidades de paisaje (SILVA ET AL, 2007). El paisaje se caracteriza por los resultados de los análisis multivariantes realizados simultáneamente con el análisis visual.

### 3. Resultados y discusión

En el Análisis Orientado a Objetos, es necesario la configuración de los niveles de escala y fusión. Para el nivel de escala se empleó un 70,5 y para el de fusión un 70 (Figura 2) y, como resultado se obtuvo un banco de datos relacional con 792 segmentos y 39 variables. Antes de procesar la ACP y AA, fueron excluidas las tres primeras columnas de la base de datos relacional; los identificadores generados para que los datos fuesen exportados directamente al programa ArcGIS. Con la base de datos relacional fueron realizados ACP y AA.

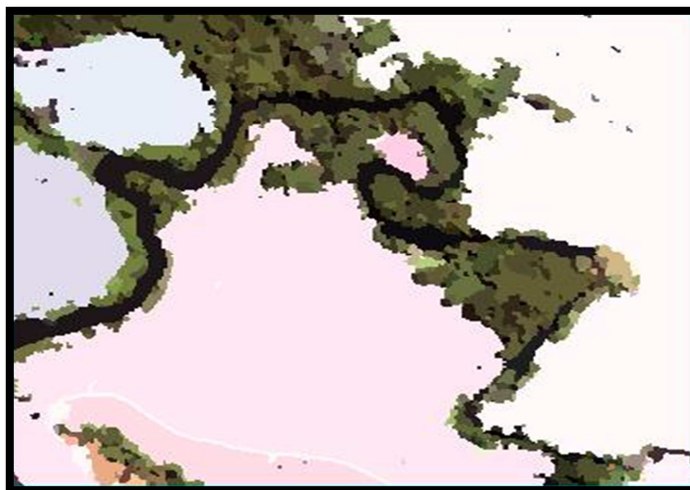


Figura 2 –Análisis Orientado a Objetos nivel de escala 70, 5 y de fusión 70.

En el primer Análisis de Componentes Principales las variables se redujeron a 22 y más del 87% de varianza se concentró en los 5 primeros componentes (Tabla 2). Se aplicó a AA y sin embargo algunas variables (dirección principal (I), elongación (G), desviación stander de los valores digitales banda 4 (AK)), desviación stander de los valores digitales banda 2 (E), se agruparon en grandes distancias euclídeas y aislados (Figura 3). Por lo tanto, se realizó un segundo análisis CP, cuando se excluyeron las variables de menor varianza (análisis de los gráficos entre componentes 1 a 5) con el objetivo de disminuir aún más la dimensionalidad de los datos (Tabla 3).

Tabla 2 – Primer Análisis de Componentes Principales

Componente Principal	Varianza (%)
1	31,72
2	23,16
3	13,15
4	11,04
5	8,07
Total	87,14
Otras variables	12,86

En el segundo ACP las variables fueron reducidas a 14 con aproximadamente el 86% de varianza en los 3 primeros componentes (Tabla 3).

Tabla 3 – Segundo Análisis de Componentes Principales

Componente Principal	Varianza (%)
1	40,54
2	23,69
3	21,42
Total	85,65
Otras variables	14,34

El Análisis de Agrupaciones se utilizó para los dos ACP, con 22 y 14 variables, figuras 3 y 4, respectivamente. A la continuación puede ser vista el análisis realizado según las variables del primer Componente Principal (Figura 3) con 4 subgrupos principales.

El primer grupo juntó 6 variables espectrales relativas a los valores de los niveles digitales de las bandas 2 y 3 en una distancia euclidiana de aproximadamente 0,15: mínimo valor digital, banda 2 (Z) promedio de los valores digitales, banda 2 (AB), máximo valor digital, banda 2 (AA), mínimo valor digital, banda 3 (AD), promedio de los valores digitales, banda 3 (AF) y máximo valor digital, banda 2 (AE). Para el segundo grupo, las correlaciones se dieron con una distancia euclidiana aproximada de 0,22 originando 2 variables espaciales: convexidad (C) y relación con agujeros (M), 2 espectrales: intensidad (Q) y matiz (O) y 1 de textura: entropía de textura (U). El tercer grupo se formó con distancia euclidiana aproximada de 0,22 y originó 3 variables espectrales correspondientes a las bandas de longitud de onda del verde y del rojo: mínimo valor digital banda 4 (AH), máximo valor digital banda 4 (AI) y promedio de los valores digitales banda 2 (AJ). En el cuarto grupo las correlaciones se dieron con distancia euclidiana aproximada de 0,22 y se formaron 4 variables espaciales que aluden a las formas del paisaje: forma rectangular (H), solidez (D), compacidad (B) y factor de forma (F).

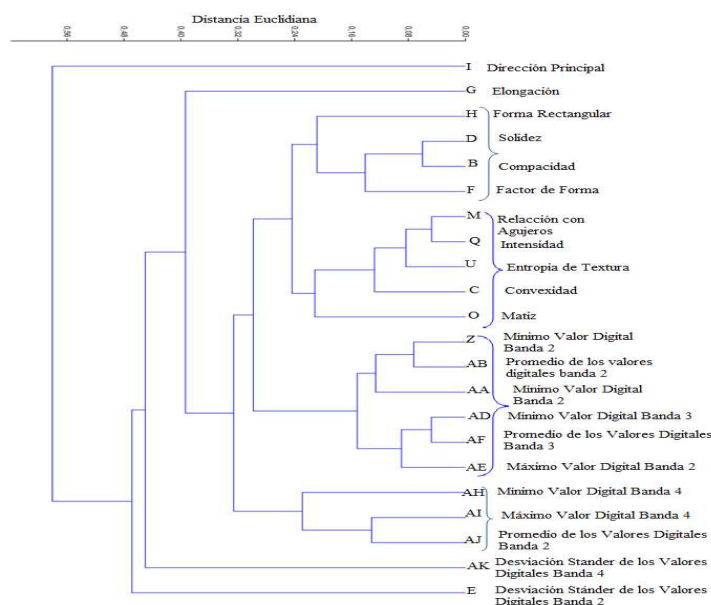


Fig. 3. Dendrograma del primer análisis de Componentes Principales con 22 variables.

La Tabla 4 muestra que la cantidad de variables espaciales, espectrales o de textura agrupadas, provenientes del primer Análisis de Componentes Principales con 22 variables.

Tabla 4 - Tipo y cantidad de variables de cada subgrupo

Análisis de Agrupaciones 1 - 22 variables		
Grupo 1		
Espacial	Espectral	Textura
6		
Grupo 2		
Espacial	Espectral	Textura
2	2	1
Grupo 3		
Espacial	Espectral	Textura
3		
Grupo 4		
Espacial	Espectral	Textura
4		
Variables con baja similitud		
Espacial	Espectral	Textura
2	2	

El Análisis de Agrupaciones realizado según las variables resultantes del segundo Componente Principal (Figura 4) generó 3 subgrupos principales. El primer subgrupo agregó 2 variables espaciales: relación con agujeros (M) y convexidad (C), 2 espectrales: intensidad (Q) y matiz (O) y 1 de textura: entropía de textura (U), agrupándose a una distancia euclidiana de aproximadamente 0,22. El segundo grupo se formó a partir de 6 variables espectrales referentes a las bandas verde y rojo: mínimo valor digital, banda 2 (Z), promedio de los valores digitales, banda 2 (AB), máximo valor digital, banda 2 (AA), mínimo valor digital, banda 3 (AD), promedio de los valores digitales, banda 3 (AF) y máximo valor digital, banda 3 (AE), en distancia euclidiana aproximada de 0,16. El tercer grupo se formó en una distancia euclidiana aproximada de 0,24, compuesto por 3 variables espectrales: mínimo valor digital banda 4 (AH), máximo valor digital, banda 4 (AI) y promedio de los valores digitales, banda 4 (AJ), todas referentes a las bandas del infrarrojo cercano.

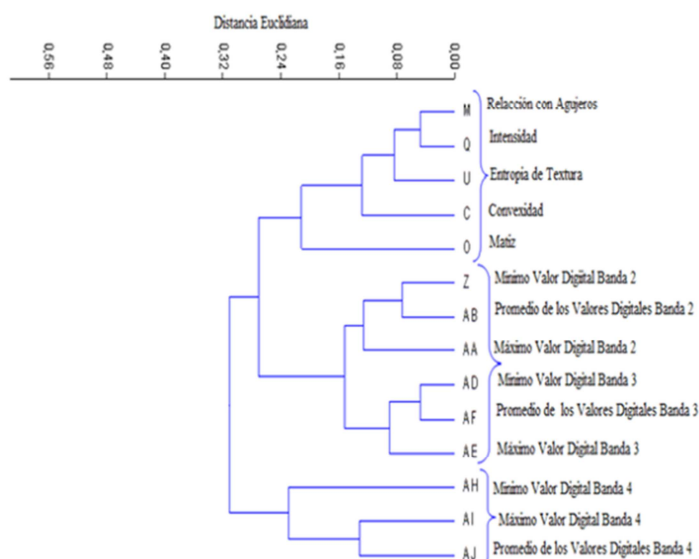


Figura 4 – Dendrograma del segundo Análisis de Componentes Principales con 14 variables.

La Tabla 5 muestra la cantidad de variables espaciales, espectrales o de textura agrupadas, provenientes del segundo Análisis de Componente Principales con 14 variables. Nótese que en este análisis no se originaron variables aisladas, siendo eliminadas aquellas con una baja correlación con la unidad de paisaje.

Tabla 4 - Tipo y cantidad de variables de cada subgrupo

Análisis de Agrupaciones 2 - 14 variables		
Grupo 1		
Espacial	Espectral	Textura
2	2	1
Grupo 2		
Espacial	Espectral	Textura
	6	0
Grupo 3		
Espacial	Espectral	Textura
	3	

#### 4. Conclusión

Los análisis multivariantes y estadísticos utilizados en el presente estudio constituyen el método adecuado puesto que se trata de una técnica investigativa y de optimización de tiempo por resultado. Se observó que, tanto en el primer análisis de agrupaciones como en el segundo (Figuras 3 y 4) las variables espectrales fueron más significativas, con un total de 11 en cada aplicación. Las variables espaciales fueron las segundas más significativas, agrupándose 6 en el primer análisis y 2 en el segundo. Las variables de textura fueron las menos numerosas, ya que solamente se originó 1 variable en el primer y segundo análisis.



Para el Análisis de Componentes Principales, la mayor similitud se dio con los componentes espectrales, debido a la reflectancia de la vegetación y los usos en torno al río Pitangui. En el Análisis de Agrupaciones esas variables se dividieron en diferentes grupos, según los números digitales de las bandas, siendo que en el verde y el rojo los niveles digitales son bajos, y en el infrarrojo cercano altos.

Por eso, la caracterización del área da lugar a las diferentes respuestas espectrales, una vez que el comportamiento espectral no siga un patrón estático de respuesta y sí refleje fracciones de energía emitida y absorbida, lo que caracteriza el objetivo en cuestión. De esa forma, la segmentación acompañó las clases de vegetación de la ribera distribuida según diferentes niveles de crecimiento y de perímetro, así como del uso por la agricultura existente en las proximidades del río.

Los componentes espaciales tuvieron una similitud intermedia, no obstante también fue significativo debido a la morfología de la vegetación y del río.

El único componente de textura que tuvo similitud en el ACP y se formó en el AA fue la entropía, debido, principalmente, a las áreas con diversos niveles digitales provenientes de la vegetación distribuida según las diversas fases de crecimiento.

### Agradecimientos

Los autores agradecen a la Coordinación de Perfeccionamiento de Personal de Nivel Superior (CAPES), la ayuda financiera y también a la Universidade Estadual de Ponta Grossa (UEPG) por la adquisición de la imagen de satélite utilizada en el presente trabajo.

### Referencias

- ANTUNES, A. F. B., 2003, Classificação de ambiente ciliar baseada em orientação a objeto em imagens de alta resolução espacial. Tese de Doutorado, Pós Graduação em Ciências Geodésicas da Universidade Federal do Paraná, Brasil.
- BOROUCHÉ, J. M., 1982, SAPORTA. G. Análise de dados, Zahar Editores, Rio de Janeiro. 116 p.
- JOHNSON, R. A.; WICHERN, D.W., 2002, Applied Multivariate Statistical Analysis. 5. ed. New Jersey: Pearson.
- MATHER, P.M., 1999, Computer processing of remotely sensed images, New York, John Wiley e Sons, Inc. 292 p.
- MINGOTI, S. A., 2005, Análise de Dados por meio de Métodos de Estatística Multivariada: uma abordagem aplicada, Belo Horizonte, Editora UFMG, 295 p.
- MULLER E., 1997, Mapping riparian vegetation along rivers: old concepts and new methods. Aquatic Botany, 58: 411-437.
- PRICHOA, C. E., 2012, Estudo em duas unidades de paisagem da bacia hidrográfica do Rio Pitangui/PR mediante estatística multivariada e análise orientada a objetos. Dissertação de Mestrado, Universidade Estadual de Ponta Grossa, Paraná, 141p.
- REGAZZI, A. J., 2001, Análise Multivariada, Viçosa, Universidade Federal de Viçosa, Centro de Ciências Exatas e Tecnológicas, Departamento de Informática. 166p.
- ROBINSON, D. J. REDDING, N. J.; CRISP, D. J. Implementation of a fast algorithm for segmenting SAR imagery, Australia, DSTO, Electronics and Surveillance Research Laboratory, DSTO-TR-1242, Jan 2002.
- SCHOWENGERDT, R. A., 1997, Remote Sensing: Models and Methods for Image Processing, 3. ed. London: Academic Press, 560 p.
- SILVA, J. dos S. et al., 2006, Espacialização da Homogeneidade e Heterogeneidade do meio biofísico para planejamento ambiental utilizando análise multivariada. In: simp. de geotecnol. no pantanal, 1, Anais... Campo Grande, Novembro.
- SILVA, J. dos S. et al., 2007, Zoneamentos ambientais: quando uma unidade territorial pode ser considerada homogênea? Revista Brasileira de Cartografia, n.59, v. 1, Abril.
- VENTURIERI, A.; SANTOS, J.R., 1998, Técnicas de classificação de imagens para análise de cobertura vegetal. In: Assad, E. D.; Sano, E. E. (Ed.). Sistema de Informações Geográficas, Brasília: Embrapa-SPI/Embrapa-CPAC, Capítulo 18, p. 351-371.
- VICINI, L.; SOUZA. A. M. ,2005, Análise Multivariada da teoria à prática. Santa Maria. UFSM. 215P.